The Leeds University maternity audit project

B. WILSON1, J. G. THORNTON, J. HEWISON2, R. J. LILFORD3, I. WATT4, D. BRAUNHOLTZ3 AND M. ROBINSON5

1Centre for Reproduction Growth and Development, 2Department of Psychology and 3Nuffield Institute, Leeds University, Leeds, 4Department of Public Health, Birmingham University, Birmingham and 5Department of Health Studies, York University, York, UK

Abstract

Objectives. To measure levels of and changes in compliance with evidence-based recommendations in obstetrics in the UK. To identify barriers to and factors associated with compliance.

Design. A quantitative case-note audit for 1988 and 1996, and a qualitative interview study of key staff.

Setting. Twenty maternity units, selected at random from all UK units

Subjects. Fifty consecutive cases of pre-term delivery (PTD), Caesarean section (CS), instrumental delivery (ID), and perineal repair (PR) operations in each period in each unit. The lead clinician, midwifery manager, a senior midwife, neonatologist, and middle-grade obstetrician in each unit.

Main outcome measures. Maternal steroid use in PTD, antibiotic use in CS, use of the ventouse (vacuum extractor) rather than forceps as instrument of first choice for ID, and use of polyglycolic acid (PGA) sutures for PR in each time period. Facilities for implementing, staff attitudes to, and the degree of planning to follow each recommendation.

Main results. The median proportion of ventouse as instrument of first choice in each unit was 8% (range 0–32%) in 1988, rising to 64% (range 0–98%) in 1996. PGA use for PR was 0% (range 0–30%) in 1988, and 72% (range 0–100%) in 1996. Steroid use for eligible PTD was median 0% (range 0–23%) in 1988, rising to 82% (range 63–95%) in 1996. Antibiotic use for CS was 7% (range 0–25%) rising to 84% (range 10–100%) in 1996. There was no relationship between unit size, type of unit, facilities, staff attitudes or degree of planning, and compliance with the recommendations, nor was the level of adherence to one standard typically correlated with adherence to the others. However, there was a positive correlation ($R = 0.6$, $P < 0.005$) between local availability of the Cochrane database of perinatal trials and unit compliance with the audit standards in the latter time period.

Conclusions. We have documented a massive shift in practice in line with the evidence, although many units still have substantial room for improvement. About 2000 wound infections, 200 deaths due to prematurity, nearly 8000 women in pain from catgut sutures, and 1500 cases of severe perineal trauma from forceps remain preventable. The reasons why units vary remain obscure, although the qualitative interviews often revealed local factors such as key enthusiastic staff. There was no sign of evidence being positively driven into practice by any systematic managerial process. The relationship between Cochrane availability and high-standard care may be simply a marker of commitment to the evidence, but it remains plausible that if senior staff make Cochrane available for their juniors, audit compliance improves.

Keywords: audit, clinical standards, evidence-based care, pregnancy

In the UK, agreement on what constitutes evidence-based care for pregnancy and childbirth is relatively advanced because of the efforts of the National Perinatal Epidemiology Unit in Oxford, UK, which has collected and disseminated evidence from randomized controlled trials since the 1980s. Systematic reviews have been published in books [1,2], and computer databases [3] are widely available through the Cochrane collaboration. These form the basis of a range of guidelines produced by the Royal College of Obstetricians and Gynaecologists (RCOG). However, the recommended practices may have been only patchily implemented. Allegedly, only one in five women in the UK received steroids prior to...
pre-term delivery less than 8 years ago [4], although it was not clear what proportion had a contraindication or insufficient time to administer them. If such claims are true of such a well-publicized recommendation, compliance with others might be even lower. Unfortunately, these data are based on small studies that did not measure eligibility adequately, and which may be out of date, and routine data are inadequate to check today’s figures precisely.

If compliance is low those responsible for quality of care should take steps to improve it. In the UK this would have to a large extent be the role of those with a responsibility for the newly introduced function of clinical governance. However, the best methods for translating evidence into practice are unclear, as evidenced by a recent review, which identified no less that 44 systematic reviews of 102 different studies of methods to do this [5]. The main conclusion was that dissemination activities by themselves were rarely effective, there were no ‘magic bullets’, and a diagnostic analysis identifying barriers to change should precede interventions to effect change. The first part of this project was to provide up-to-date estimates of rates of compliance with evidence-based recommendations and to measure changes over time. The second part of the project comprises such a diagnostic analysis, albeit undertaken after many of the recommendations had been circulated.

Methods

Four audit standards underwritten by evidence-based recommendations were selected for study.

1. For perineal injury, polyglycolic acid sutures (Dexon or Vicryl) should be used for repair of both the deep layers and skin.
2. All women undergoing Caesarean section should receive prophylactic antibiotics.
3. All women expected to deliver pre-term (<34 weeks [6]) should be administered corticosteroids.
4. The ventouse (a vacuum cup attached to the baby’s head) should be the instrument of first choice for operative vaginal delivery, in preference to the obstetric forceps.

These topics were selected largely for practical reasons, namely that cases could be easily ascertained from the statutory labour-ward record book. Some topics such as use of postnatal anti-D immunoglobulin for Rhesus prophylaxis were excluded because compliance was already documented to be very high [6]. Others, such as use of external cephalic version for breech presentation or the offer of induction post-term, would have required review of all or most records because few units keep a computer or paper record of cases with mal-presentation or post-maturity, and would therefore have been very expensive. We also intended to examine one further audit standard, which emerged after submission of the protocol, namely that all women with eclampsia should be treated with magnesium sulphate. However, eclampsia is so uncommon that only a few cases could be studied opportunistically, and results demonstrating a massive and rapid change of practice have already been reported elsewhere [7].

Twenty maternity units were selected at random from a full list of those in England and Wales held by the RCOG. All hospitals initially selected agreed to participate. Units were classified as teaching and non-teaching hospitals, and their number of annual deliveries was recorded.

We measured compliance with each audit standard for two periods, the years 1988 and 1996. The study began during 1998, while 1996 was the most recent year for which delivery records were unlikely to still be in use. 1988 was the latest year before the randomized trial evidence became widely available to obstetricians with publication of the book Effective Care in Pregnancy and Childbirth in 1989 [1]. Although individuals may have been aware of the evidence before that date, and some units may have achieved high compliance by accident, no systematic efforts to disseminate evidence had been made at that time. The actual evidence, in terms of published randomized controlled trials, on which the RCOG recommendations were based did not alter substantially between 1988 and 1996 for any of the standards studied. For example, although there was evidence for the effectiveness of steroids as early as 1972 [8] many review articles and textbooks regarded this as inconclusive until Crowley’s review in 1990 [9]. The RCOG promulgated national guidelines in 1992.

We identified an audit clerk in each unit and invited them to Leeds for 2 days of training. The purpose of the study and the clinical justification for each audit standard were explained. Each clerk brought two sets of notes for each topic from their own hospital, for each of which they and another clerk independently completed an audit data form. The results were compared and checked by two of the authors (JGT or RJL) and the form design was modified in response to advice from the clerks about availability and accuracy of local data. Each clerk then completed five finalized audit forms from standard notes for each audit standard, which were checked before starting the project.

Each clerk identified 50 sets of records of Caesarean deliveries, deliveries before 34 completed weeks, and operative vaginal deliveries from each time period from the delivery register. The latter 50 records were used to assess both the instrument of first choice for the operative delivery and the suture material used for perineal repair. Suitable records were identified by simultaneously searching forward and backwards from the first of June in each index year until 50 cases were identified.

For the ventouse, suture material, and antibiotic use at Caesarean section audit, the clerks were able to classify records unambiguously with little difficulty. Any record of administration at the appropriate time was regarded as evidence that antibiotics or steroids had been given, even if a drug prescription chart could not be found. A more elaborate system was required to assess whether patients who had delivered before 34 weeks had been eligible for steroids and whether there had been, in prospect, sufficient time for them to act. If steroids had been given at all, the cases was classed as ‘audit standard met’. Otherwise patients either admitted
with a diagnosis of pre-term labour for more than 3 hours or delivered electively were classed as 'eligible and not given steroids'. This is a conservative algorithm in that some people in whom delivery appears imminent may not deliver as soon as expected. Patients admitted with a diagnosis of pre-term labour less than 3 hours before delivery or in whom pre-term labour was never diagnosed were classed as 'not eligible for steroids'. For example, a woman admitted with abdominal pain and a closed cervix, with a reasonable diagnosis of urinary infection or non-specific pain, who nevertheless went on to deliver precipitously would be classed as ineligible. All the latter cases, and a one in 10 subset of those in whom the audit standard was met, were reviewed independently by one of three experienced clinicians. Any disagreements were reviewed again by one of the authors (JGT) and a final classification was made.

The result was a level of compliance with each of the four audit standards for each unit at each time period. Finally, a hospital level of audit compliance for each time period was calculated as the mean of each of these four standards.

The research assistant interviewed five people from each unit [the medical director or senior obstetric consultant, the unit manager (who may have been a midwife or administrator), the paediatrician with most responsibility for neonatal care, a clinically active midwife, and a middle grade obstetrician]. The aim was to measure the degree to which respondents had moved along the continuum of the 'theory of implementation intentions' [10,11]. This suggests that behavioural change can be divided into two phases, an intention/motivation phase and an implementation phase.

The interview was divided into two parts. In the first part, respondents were asked if there was a unit policy for each topic, and the responses classified as no policy, unclear if there is a policy, unclear if the policy follows the guidelines, policy differs from the guidelines, or policy follows the guidelines. The following questions concerned respondent's knowledge of and attitudes towards evidence-based practice and the Cochrane collaboration, and towards the four specific study guidelines. Information on the local availability of the Cochrane database was also collected. If respondents were unfamiliar with the content of a recommendation, that in the analyses reported below. The correlation between scores (improvement in compliance) were highly correlated with final scores, only the latter are used as outcome measures in the analyses reported below. The correlation between scores for different standards varied. For example, in 1996, there was a weak positive correlation between compliance with the audit standards for instrumental delivery and urinary infection or non-specific pain, who nevertheless went on to deliver precipitously would be classed as ineligible. All the latter cases, and a one in 10 subset of those in whom the audit standard was met, were reviewed independently by one of three experienced clinicians. Any disagreements were reviewed again by one of the authors (JGT) and a final classification was made.

The second phase of the interview covered the extent to which implementation had actually occurred. Respondents were asked if any explicit attempts to change practice had been made, and if so, what these had been. Had any guidelines been written, had any formal attempt been made at dissemination, or had any co-ordinated action to implement the unit policy been taken? At unit level, having an explicit policy was taken as evidence of intention to follow a recommendation.

The interviews were audiotaped, transcribed, and coded using standard methods of content analysis employed in social surveys [12]. For each question, data from the full sample were used to devise the coding frame, and individual responses classified accordingly. Numeric codes were then assigned to the classified data for the purposes of quantitative analysis. The audit and interview results were analysed using SPSS. The outcome variable was always compliance with the audit standard in the second time period (1996), either for each standard individually or aggregated by unit as appropriate. First the relationship of teaching/non teaching and the paediatrician with most responsibility for neonatal care, a clinically active midwife, and a middle grade obstetrician]. The aim was to measure the degree to which respondents had moved along the continuum of the 'theory of implementation intentions' [10,11]. This suggests that behavioural change can be divided into two phases, an intention/motivation phase and an implementation phase.

The interview was divided into two parts. In the first part, respondents were asked if there was a unit policy for each topic, and the responses classified as no policy, unclear if there is a policy, unclear if the policy follows the guidelines, policy differs from the guidelines, or policy follows the guidelines. The following questions concerned respondent's knowledge of and attitudes towards evidence-based practice and the Cochrane collaboration, and towards the four specific study guidelines. Information on the local availability of the Cochrane database was also collected. If respondents were unfamiliar with the content of a recommendation, that in the analyses reported below. The correlation between scores (improvement in compliance) were highly correlated with final scores, only the latter are used as outcome measures in the analyses reported below. The correlation between scores for different standards varied. For example, in 1996, there was a weak positive correlation between compliance with the audit standards for instrumental delivery and steroid use ($R = 0.46$, $P = 0.04$, two tailed) and between the perineal suture and antibiotic standard for the same time period ($R = 0.51$, $P = 0.02$). Neither of these was predicted in advance. Otherwise, there were no other significant correlations among 14 comparisons, and the above significance tests would be rendered non-significant if an appropriate adjustment were made for multiple comparisons. There was no difference between the 1996 compliance rates for teaching or non-teaching hospitals ($P = 0.97$), and no relationship between compliance and the size of the hospital ($P = 0.32$).

There were 88 taped interviews. In the latter time period (1996) only six units had half or more respondents reporting full access to the Cochrane database. There was a positive relation between average unit compliance with the four audit
Figure 1 The rate of compliance with the steroid use standard in each unit for each time period.

Figure 2 The rate of compliance with the instrumental delivery standard in each unit for each time period.

Figure 3 The rate of compliance with the antibiotic standard in each unit for each time period.

Figure 4 The rate of compliance with the suture material standard in each unit for each time period.

limited access (23/59). The trend was, if anything, in the opposite direction. Staff seniority may explain these patterns. Access was higher in senior than junior staff (20/39 versus 9/49; \( P = 0.002 \)), but knowledge was lower (8/39 versus 21/49; \( P = 0.047 \)).

Senior and junior staff were just as likely to have positive attitudes (28/39 versus 36/49), but in senior staff, the majority (18/28) of those with positive attitudes had full Cochrane access, whereas this was true for only a minority of more junior staff (9/36; \( P = 0.004 \)). The two people with negative attitudes who had full Cochrane access were also both seniors.

It was possible to obtain information on awareness of specific recommendation content, and the source of that awareness, in 80 of the 88 interviews. Some managers, for example, did not have a clinical background, and preferred not to comment on what they perceived as clinical matters, and some paediatricians preferred not to comment on the antibiotic use and suture material standards. Awareness was high for the steroid recommendation (79/80), but far from complete for the ventouse (66), suture (58) and antibiotics (36) recommendations. Respondents reported having heard about the different recommendations from a variety of sources. Nine people mentioned the RCOG in connection
with the ventouse recommendation, and six, two, and three in connection with the steroid, suture, and antibiotic recommendations, respectively. The numbers of individuals mentioning directorate/departmental sources were three, five, five, and four, for the same four standards, respectively. Other sources of information included preparation for examinations, clinical practice, the research literature, and audit meetings.

Expressed attitudes to the individual recommendations were generally favourable: antibiotics 49 favourable versus four unfavourable, steroids 84 versus two, sutures 61 versus three, and ventouse 58 versus nine. There was no correlation between unit level respondents average attitude (negative mark for unfavourable attitude) and compliance with each audit standard (instrumental delivery $R = 0.023$, perineal suture material $R = 0.13$, antibiotics $R = 0.32$, steroid use $R = 0.29$). The apparently favourable attitudes to the ventouse were accompanied by qualifying remarks in 27/58 care in the UK NHS.

The proportion of respondents reporting that their unit had written guidelines in accord with the standards varied by standard (steroids 72/88; ventouse 12/70 plus eight reporting a guideline differing from the audit standard; suture material 35/70 plus one guideline differing from the standard; antibiotics 42/70 plus one differing from the standard. Note that paediatricians were only asked about the steroid recommendation). Units with higher proportions of respondents reporting suture material guidelines also had higher compliance with the suture material standard ($\rho = 0.47$, $P = 0.035$), but for the other standards there was no relation between having written guidelines and compliance with audit standards.

There was little evidence of systematic planning to implement any recommendations in any units. Relatively few units had made any explicit attempt to disseminate the policies, or designed an implementation strategy and facilitated adoption of the policy by, for example, sending people on training courses, buying new equipment, or ensuring that only the appropriate materials were available. The figures were:

1. Ventouse – 1/20 units had disseminated guidelines but none had taken co-ordinated managerial action.
2. Steroids – 7/20 disseminated and none actioned further.

In one of the four sites where a suture policy had been actioned, only a minority of respondents believed that their unit had a suture policy, so there were only three instances out of a possible 80 where a successfully disseminated policy and an action plan occurred together. None of the actions or lack thereof correlated with recommendation compliance, but the numbers were small.

**Discussion**

We have shown a dramatic rise in adherence to the four evidence-based recommendations over the 8 years since 1988. It is not possible to say how much of this resulted directly from the assembly of the evidence by the Cochrane collaboration and its forerunners, and how much from various dissemination activities such as the RCOG audit guidelines or the National Health Service (NHS) audit programme. Nevertheless, it is clear that over a relatively short time period, obstetricians and midwives have altered their practice in response to evidence. It is no longer possible to claim that only 20% of eligible women are receiving steroids. However, adherence rates are still below 100% in many units, and in some units considerably below this level. As a result, large numbers of women and babies are receiving substandard care in the UK NHS.

This improvement in adherence to recommendations is despite the fact that few units have access to the Cochrane database, have prepared or disseminated guidelines, or have taken any active steps to implement recommendations. The explanation for the range of unit compliance levels remains unexplained. With one exception, none of the knowledge attitudinal, or behavioural characteristics, which we recorded for each unit, explained the difference. The exception was access to the Cochrane database. This may mean that access to the database is causing high compliance, but it is more likely that access is a marker of a type of staff or organizational characteristic, which goes with the following of evidence-based recommendations. A plausible interpretation of these data is that senior people with positive attitudes to Cochrane arrange access to the database for themselves and, to a lesser extent, for their staff. At unit level, there were no sites in which junior people had access but senior people did not, so essentially, the units in which a greater proportion of staff had full Cochrane access were the ones in which access was available to some junior as well as senior staff.

Essentially, senior staff arranging access to Cochrane for their juniors correlates with high levels of compliance. This is plausibly a causative relation.

The shortfall in compliance with the recommendations in the latter time period is all the more important because we took considerable care to ensure that legitimate reasons for non-compliance such as admission in advanced labour, were excluded. We also ensured that we classified the choice of instrument as correct if the ventouse was used as first choice but delivery was completed with another instrument. This means that any residual shortfall is likely to be genuine, although our algorithm on steroid use is conservative. Our algorithm would have underestimated the steroid administration shortfall to eligible women since those eligible women in whom steroids were omitted but who did not go on to deliver prematurely, would not be classed as a failure to adhere to the standard.

The specific shortfall in steroid use after legitimate reasons for non-prescription has been identified as similar to that seen in thrombolytic therapy after acute myocardial infarction. In Europe only 36% of such patients receive thrombolysis, but
after those with clinical contraindications, uncertain diagnosis, and uncertain event timing have been excluded, the shortfall in prescribing falls from 64 to 20% [13].

Nevertheless, if we assume that shortfalls in compliance in the units we studied are similar to those in other units, we can estimate the avoidable morbidity caused by failing to follow the evidence. For example, in 1996 only 72% of women were benefiting from the 60% reduction in wound infection from prophylactic antibiotics, which would imply about 2000 avoidable infections per year (assuming a 15% CS rate and 6% infection rate [14]). Similarly, we estimate that only 81% of babies who could benefit from steroids were receiving them, and can assume that 3% births occur at less than 35 weeks, of which 77% would be eligible for steroids, with a 44% rate of RDS, and 18% mortality, reducible by 50% and 40%, respectively [15]. This would imply approximately 500 avoidable cases of RDS and 200 avoidable deaths from prematurity each year in the UK. About 350 000 women require perineal repair each year, of which 25% experience short-term pain when sutured with catgut [16]. Since this can be reduced by 30% if PGA sutures are used [17] and only 61% of perineal repairs use this, we estimate that over 10 000 women experience avoidable perineal pain every year. There are about 50 000 instrumental deliveries per year in the UK, which if the forceps were used for all would be associated with 10 000 unnecessary cases of severe pain and 5000 of severe perineal trauma [18]. Since even in 1996 the ventouse was used only 67% of the time, this implies 3000 unnecessary cases of severe pain and 1500 cases of severe trauma. These estimates are similar to the rates of forceps (5.8%) and ventouse delivery (4.8%) recorded in national statistics for 1994–1995, although those figures refer to instrument used to achieve delivery rather than of first choice [19]. Nevertheless, this is a considerable burden of avoidable suffering.

The varying response to different evidence-based recommendations may reflect different levels of belief in the evidence from the randomized controlled trials. This has been documented to vary widely [7]. In general, respondents to our survey were in agreement with the recommendations. What our results appear to show is that, contrary to much prevailing opinion, doing good research and disseminating it does result in a change of practice, provided that one is prepared to wait for the necessary consensus. It remains to be seen whether national organizations (such as the National Institute for Clinical Excellence in the UK) will carry greater authority than professional organizations, although they should continue to work with them. The evidence would suggest that a systematic and managed approach to the uptake of research findings was rudimentary over the time period of the study. The Department of Health has placed a duty on all managers to implement such as strategy forthwith. If this is effective we would predict that future research will show strong correlation between adherence to one standard and another, and interviews with staff should unmask such a systematic approach. In the meantime, it is worth reflecting that uptake of evidence now appears to be a far more rapid than in 1992 when Antman published his famous article on the inordinate delay in responding to clear-cut research results [20]. Our study shows that clinicians do respond to the evidence, albeit imperfectly, and in the case of magnesium treatment for eclampsia did so within a year [7]. Furthermore, they know the evidence, and by and large accept the results of well-conducted research. Clearly the culture, at least within the UK, has changed radically. Whether there has been a similar response to the evidence in other countries must await similar local audits.

References


Accepted for publication 12 February 2002